Citation:

López Osornio MM, Hough G, Salvador A, Chambers IV E, McGraw S, Fiszman S. Beef's optimum internal cooking temperature as seen by consumers from different countries using survival analysis statistics. *Food Quality and Preference*. 2008 Jan, 19 (1): 12-20.

Study Design:

Cross-sectional study

Class:

D - Click here for explanation of classification scheme.

Research Design and Implementation Rating:

NEUTRAL: See Research Design and Implementation Criteria Checklist below.

Research Purpose:

The purpose of this work was to predict the optimum cooking temperatures of beef based on acceptance or rejection data obtained from consumers using survival analysis statistics. Consumers from different countries, age groups and stated preference for degree of doneness were considered.

Inclusion Criteria:

Consumers who consumed cooked beef at least once a week in Argentina, USA and Spain.

Exclusion Criteria:

None reported

Description of Study Protocol:

Recruitment

Consumers who consumed cooked beef at least once a week were recruited from the cities of Nueve de Julio (Buenos Aires, Argentina), Manhattan (Kansas, USA) and Valencia (Spain).

Design

- In the USA, a Beef Steak Color Guide (American Meat Science Association, Chicago, USA) illustrates these color changes from ‘‘very rare” to ‘‘well–well done” beef. This Color Guide has six pictures of beef steaks labeled as: ‘‘55 °C -very rare”, ‘‘60°C -rare”, ‘‘63°C -medium rare”, ‘‘71°C -medium”, ‘‘77°C -well done” and ‘‘82°C -well well done”. In other countries, such as Argentine orSpain there are no local equivalents to the Color Guide.
- To allow the test to be answered by several consumers simultaneously, a total of 30 Color Guides were acquired from the National Cattlemen's Beef Association (Centennial, CO,

USA) by the Argentine group who received them by post. They were checked for similarity and sorted randomly into three groups of 10, one for each country. From Argentina they were sent to Spain and the USA.

- The pictures (without any description) from each color guide were cut out and coded with three-digit random numbers. Consumers received the six pictures monadically order of presentation was balanced over consumers.
- For each picture consumers had to tick a box indicating if they considered the meat in the picture ‘‘under-cooked”, ‘‘ok” or ‘‘over-cooked”. Pictures were evaluated under natural daylight conditions
- After evaluating the six pictures, consumers in Argentina and Spain were asked: ‘‘At home or in a restaurant, how do you normally consume beef? Rare, medium, well done or other?” The term ‘other’ was used should the first three options not satisfy the consumers’ opinion. Answers to ‘other’ were few and they were re-classified in the other three options reclassification was made considering if consumers’ opinion using others words, could be related with the other three options (rare-medium-well done) i.e. ‘‘leaning toward rare” was re-classified as ‘‘rare”.
- In the USA consumers are used to using more categories so they were asked: ‘‘At home or in a restaurant, how do you normally consume beef? Rare, medium-rare, medium, medium-well, well-done, or other?” In order to be able to compare data between countries, USA consumers’ data were re-categorized: ‘‘rare” and ‘‘medium-rare” were considered ‘‘rare”; and ‘‘medium-well” and ‘‘well-done” were considered ‘‘well-done”.
- Before leaving consumers received a reward for their participation.

Blinding

None

Intervention

None

Statistical Analysis

- Survival analysis was conducted to determine the optimum internal cooking temperature (ICT).
- In seeking beef’s optimum ICT, the acceptance and rejection patterns of consumers receiving these pictures was explained:
    - Rejection of the beef because it is under-cooked: If a consumer rejects the 63°C picture (medium rare) because it is under-cooked and accepts the 77 °C picture because it is ok, the exact ICT at which the consumer passes from ‘‘rejection because it is under-cooked” to ‘‘acceptance because it is ok” could be any value between 63 and 77°C. This is defined as interval censoring
    - Left censoring occurred if a consumer found the 55°C picture ok, thus the event of passing from under-cooked to ok would have occurred at some unknown ICT <55°C
    - If the consumer rejects all pictures because they are under-cooked, acceptance would occur for an ICT >82°C and this consumer’s data is right censored
- Rejection of the beef because it is over-cooked: If a consumer accepts the 63°C picture and rejects the 77°C picture because it is over-cooked, the exact ICT for which the consumer passes from accepting the beef to rejecting it because it is over-cooked could be any value between 63 and 77°C, a case of interval censoring. Left and right censoring can also occur for rejection of the beef because it is over-cooked.

- The likelihood function, which is used to estimate the failure function, is the joint probability of the given observations of the N consumers. In this study there were two likelihood functions: Lu (undercooked) and Lo (over-cooked).
- The parameters of the log-linear model are obtained by maximizing the likelihood functions. The likelihood function is a mathematical expression which describes the joint probability of obtaining the data actually observed on the subjects in the study as a function of the unknown parameters of the model being considered.
- In order to establish whether consumer age group, stated preference for degree of doneness and country of residence influenced rejection times, a log-linear regression model with inclusion of covariates was applied
- Once the likelihood is formed for a given model, specialized software can be used to estimate the parameters (b coefficients and r) that maximize the likelihood function for the given experimental data
- The CensorReg procedure from S-PLUS (Insightful Corporation, Seattle, USA) was used to estimate the models' parameters, quantiles and corresponding standard deviations
- Five percent or less was considered for significance.

Data Collection Summary:

Timing of measurements

Not applicable, cross-sectional survey.

Dependent variables

Not applicable

Independent variables

Not applicable

Control variables

Not applicable

Description of Actual Data Sample:

- ***Initial N:*** 102 were recruited in each city (total 306)
- ***Attrition (final N):*** 102 per city
- ***Age:***
  - Half were aged ranging from 21 to 30 years (young)
  - Half were aged ranging from 40 to 60 years (middle-aged)
- ***Ethnicity:*** Argentine, Spanish, and USA.
- ***Other relevant demographics:*** Approximately half were female and half male
- ***Anthropometrics:*** None specified
- ***Location:*** Nueve de Julio (Buenos Aires, Argentina), Manhattan (Kansas, USA) and Valencia (Spain).

Summary of Results:

- Data for six consumers that illustrate the interpretation given to each subject's data:
    - Consumer one rejected the first, accepted pictures of intermediate degrees of doneness and rejected those that were over-cooked. The exact Internal Cooking Temperature (ICT) below which this consumer rejects steaks because they were under-cooked is unknown; it was between 63°C and 71°C, and his data were thus interval censored for under-cooked rejection. Analogously, the exact ICT above which this consumer rejected the steaks because they were over-cooked is unknown; it was between 71°C and 77°C, and his data were interval censored for over-cooked rejection.
    - Consumer two rejected all steaks because they were under-cooked, not finding any sample ok. Thus his data were right censored for under- and over-cooked rejection
    - Consumer three behaved similarly to consumer one, only rejecting the first ICT and those cooked at 77°C and 82°C
    - Consumer four rejected the first three ICT's and then found the rest ok. His data were interval censored for under-cooked rejection and right censored for over-cooked rejection
    - Consumer five presented inconsistency
    - Consumer six was so inconsistent that his data had to be removed
- 18 consumers presented this type of data (one from Argentina, 12 from Spain and five from USA)
- Subjects who indicated they had vision problems were excluded from the test, there could have been some consumers who did not report their problem and thus produced inconsistent results
- The magnitude of the coefficients indicates that the factor most influencing rejection due to being under-cooked was stated preference, followed by country of residence and age group
- Also, stated preference influenced rejection due to under-cooked more than rejection due to over-cooked
- Age, on the other hand, influenced rejection due to over-cooked more than rejection due to under-cooked
- These parameters can be used to relate percent consumer rejection to ICT for each event
- Argentina had the highest percentage who ask for beef ''well-done", while USA had the highest percentage who ask for beef ''rare"; Spain had the highest percentage who ask for beef cooked to ''medium"
- As expected, for a given ICT, for example 71°C, the probability of rejection due to under-cooked is higher for those consumers that stated they prefer ''well done" beef. For a given ICT, for example 81°C, the probability of rejection due to over-cooked is higher for those consumers that stated they prefer their beef ''rare".
- The 55°C picture was rejected as under-cooked by almost all consumers, including those who stated they preferred ''rare" beef. At the other extreme, the 82°C picture was rejected as under-cooked by 29% of those consumers who stated they preferred their beef ''well-done".
- Not all consumers found the 82°C picture over-cooked; in fact, 65% of those who stated they preferred ''rare" beef found this picture over-cooked, leaving a substantial 35% who did not find this picture overcooked. These observations are also valid for Spanish and USA consumers. The AMSA Color Guide would have to be revised to shift the ICT range to higher values.
- For the over-cooked event there is a single curve as country effect was not significant
- For a given ICT, for example 75°C, rejection probability due to the beef being under-cooked was 63%, 60% and 56% for consumers from Argentine, Spain and USA, respectively. The magnitude of country differences was small. Although sampling 102 consumers from one city is not representative of a country's whole population, the small differences found

between the cities suggests that similar results would have been obtained from a larger consumer sample in each country.
- For the under-cooked event both age groups practically overlapped
- For a given ICT, for example 82°C, the middle-aged consumers tended to have lower rejection probability (16%) than the younger consumers (23%) due to the beef being over-cooked
- As country (Argentina, Spain and USA), age (young and middle-aged) and stated preference (rare, medium and well-done) were significant for the under-cooked event, a total of 18 optimum ICTs were calculated, one for each covariate combination.

Author Conclusion:

- Survival analysis methodology was an adequate tool to estimate an optimum ICT of beef, based on consumer acceptability
- The experimental sensory work applied in this study was relatively simple; only 102 consumers in a country had to decide if they found successive samples under-cooked, ok or over-cooked
- Stated preference of degree of doneness was the covariate that most influenced under-cooked and over-cooked rejection probabilities (ITC)
- The influences of country of residence and age were of smaller magnitude
- The number of consumers who were not satisfied with the calculated optimum ICTs was relatively high. This means that when consumers say they want their steaks cooked to a determined degree of doneness it is difficult to understand what many of them really mean. Because the optimum ICTs were high, revisions to the MSA Color Guide should be considered to cover a higher ICT range.

Reviewer Comments:

- *The clinical relevance of this study is unclear, as it only deals with consumer preference and does not account for food safety*
- *It is unclear how participants were recruited and what their characteristics were, other than the fact that they ate beef more than once per week. Recruitment strategies would be useful in order to determine if participants were a representative sample of the population.*
- *Although the clinical relevance of evaluating consumer preference is not clearly established in this study, the conclusions that the AMSA Color Guide may need adjustments are valid.*

Research Design and Implementation Criteria Checklist: Primary Research

Relevance Questions		
1.	Would implementing the studied intervention or procedure (if found successful) result in improved outcomes for the patients/clients/population group? (Not Applicable for some epidemiological studies)	N/A
2.	Did the authors study an outcome (dependent variable) or topic that the patients/clients/population group would care about?	Yes

3.	Is the focus of the intervention or procedure (independent variable) or topic of study a common issue of concern to nutrition or dietetics practice?	Yes
4.	Is the intervention or procedure feasible? (NA for some epidemiological studies)	N/A

Validity Questions

1.	**Was the research question clearly stated?**	Yes
1.1.	Was (were) the specific intervention(s) or procedure(s) [independent variable(s)] identified?	Yes
1.2.	Was (were) the outcome(s) [dependent variable(s)] clearly indicated?	Yes
1.3.	Were the target population and setting specified?	Yes
2.	**Was the selection of study subjects/patients free from bias?**	No
2.1.	Were inclusion/exclusion criteria specified (e.g., risk, point in disease progression, diagnostic or prognosis criteria), and with sufficient detail and without omitting criteria critical to the study?	No
2.2.	Were criteria applied equally to all study groups?	???
2.3.	Were health, demographics, and other characteristics of subjects described?	No
2.4.	Were the subjects/patients a representative sample of the relevant population?	No
3.	**Were study groups comparable?**	N/A
3.1.	Was the method of assigning subjects/patients to groups described and unbiased? (Method of randomization identified if RCT)	N/A
3.2.	Were distribution of disease status, prognostic factors, and other factors (e.g., demographics) similar across study groups at baseline?	N/A
3.3.	Were concurrent controls used? (Concurrent preferred over historical controls.)	N/A
3.4.	If cohort study or cross-sectional study, were groups comparable on important confounding factors and/or were preexisting differences accounted for by using appropriate adjustments in statistical analysis?	N/A
3.5.	If case control or cross-sectional study, were potential confounding factors comparable for cases and controls? (If case series or trial with subjects serving as own control, this criterion is not applicable. Criterion may not be applicable in some cross-sectional studies.)	N/A

3.6.	If diagnostic test, was there an independent blind comparison with an appropriate reference standard (e.g., "gold standard")?	N/A
4.	**Was method of handling withdrawals described?**	???
4.1.	Were follow-up methods described and the same for all groups?	???
4.2.	Was the number, characteristics of withdrawals (i.e., dropouts, lost to follow up, attrition rate) and/or response rate (cross-sectional studies) described for each group? (Follow up goal for a strong study is 80%.)	No
4.3.	Were all enrolled subjects/patients (in the original sample) accounted for?	???
4.4.	Were reasons for withdrawals similar across groups?	N/A
4.5.	If diagnostic test, was decision to perform reference test not dependent on results of test under study?	N/A
5.	**Was blinding used to prevent introduction of bias?**	No
5.1.	In intervention study, were subjects, clinicians/practitioners, and investigators blinded to treatment group, as appropriate?	N/A
5.2.	Were data collectors blinded for outcomes assessment? (If outcome is measured using an objective test, such as a lab value, this criterion is assumed to be met.)	N/A
5.3.	In cohort study or cross-sectional study, were measurements of outcomes and risk factors blinded?	No
5.4.	In case control study, was case definition explicit and case ascertainment not influenced by exposure status?	N/A
5.5.	In diagnostic study, were test results blinded to patient history and other test results?	N/A
6.	**Were intervention/therapeutic regimens/exposure factor or procedure and any comparison(s) described in detail? Were interveningfactors described?**	N/A
6.1.	In RCT or other intervention trial, were protocols described for all regimens studied?	N/A
6.2.	In observational study, were interventions, study settings, and clinicians/provider described?	N/A
6.3.	Was the intensity and duration of the intervention or exposure factor sufficient to produce a meaningful effect?	N/A
6.4.	Was the amount of exposure and, if relevant, subject/patient compliance measured?	N/A
6.5.	Were co-interventions (e.g., ancillary treatments, other therapies) described?	N/A
6.6.	Were extra or unplanned treatments described?	N/A

6.7.	Was the information for 6.4, 6.5, and 6.6 assessed the same way for all groups?	N/A
6.8.	In diagnostic study, were details of test administration and replication sufficient?	N/A
7.	**Were outcomes clearly defined and the measurements valid and reliable?**	???
7.1.	Were primary and secondary endpoints described and relevant to the question?	Yes
7.2.	Were nutrition measures appropriate to question and outcomes of concern?	Yes
7.3.	Was the period of follow-up long enough for important outcome(s) to occur?	N/A
7.4.	Were the observations and measurements based on standard, valid, and reliable data collection instruments/tests/procedures?	???
7.5.	Was the measurement of effect at an appropriate level of precision?	???
7.6.	Were other factors accounted for (measured) that could affect outcomes?	???
7.7.	Were the measurements conducted consistently across groups?	N/A
8.	**Was the statistical analysis appropriate for the study design and type of outcome indicators?**	Yes
8.1.	Were statistical analyses adequately described and the results reported appropriately?	Yes
8.2.	Were correct statistical tests used and assumptions of test not violated?	Yes
8.3.	Were statistics reported with levels of significance and/or confidence intervals?	Yes
8.4.	Was "intent to treat" analysis of outcomes done (and as appropriate, was there an analysis of outcomes for those maximally exposed or a dose-response analysis)?	N/A
8.5.	Were adequate adjustments made for effects of confounding factors that might have affected the outcomes (e.g., multivariate analyses)?	Yes
8.6.	Was clinical significance as well as statistical significance reported?	Yes
8.7.	If negative findings, was a power calculation reported to address type 2 error?	N/A
9.	**Are conclusions supported by results with biases and limitations taken into consideration?**	Yes
9.1.	Is there a discussion of findings?	Yes
9.2.	Are biases and study limitations identified and discussed?	Yes
10.	**Is bias due to study's funding or sponsorship unlikely?**	???

10.1.	Were sources of funding and investigators' affiliations described?	???
10.2.	Was the study free from apparent conflict of interest?	???